

WHAT IS CLAIMED IS:

1. A method for Uniform Resource Locator (URL) filtering, comprising:
receiving an event notification upon the occurrence of an event associated
with a received URL;

5 searching, in response to said event notification, a lexical search tree data
structure storing a plurality of URLs for said received URL; and

processing said received URL in response to said received URL not matching
any of said plurality of URLs stored in said lexical search tree data structure.

10 2. The method of claim 1, wherein said searching said lexical search tree
comprises:

determining a hash value for said received URL;

determining a branch associated with a root node of said lexical search tree
data structure corresponding to said hash value, said branch along with said root node
15 representing at least one URL of said plurality of URLs, said branch having one or
more leaf nodes linked hierarchically to one another, each leaf node representing an
element of said at least one URL; and

traversing only said branch to find a match between said at least one URL and
said received URL.

20 3. The method of claim 2, wherein said determining a hash value
comprises:

determining a first element of said received URL; and

determining a hash value for said first element.

25 4. The method of claim 3, said hash value being the ASCII code for said
first element.

5 5. The method of claim 2, wherein said traversing only said branch comprises comparing successive elements of said received URL with successive elements of said at least one URL stored in successive leaf nodes of said one or more leaf nodes so long as said successive elements of said received URL match said successive elements of said at least one URL.

6. The method of claim 2, wherein said traversing only said branch further comprises:

10 determining a twig associated with said branch at a point of divergence between said at least one URL and said received URL, said twig representing a terminating substring of a second URL of said plurality of URLs; and

traversing said twig to find a match between a terminating substring of said received URL and said terminating substring represented by said twig.

15 7. The method of claim 6, wherein said traversing said twig comprises comparing successive elements of said terminating substring of said received URL with successive elements of said terminating substring of said second URL represented by said twig so long as said successive elements match.

20 8. The method of claim 5, wherein said traversing only said branch further comprises:

setting a current node pointer to point to a leaf node of said one or more leaf nodes;

25 setting a target signature pointer to point to an element of said received URL;

in response to a value of said leaf node pointed to by said current node pointer being equal to a wild card character and a value of the element pointed to by said target signature pointer being equal to a value of the next leaf node following the leaf node pointed to by said current node pointer, updating said current node pointer to point to a leaf node following said next leaf node.

9. The method of claim 1, wherein said receiving said event notification comprises receiving said event notification upon the occurrence of an event selected from the group consisting of a URL map event and a receive raw data event.

5 10. A system for Uniform Resource Locator (URL) filtering, comprising:
a web server operable to receive a URL request from a client; and
a filter operable, upon receiving an event notification relating to said URL request from said web server, to search a lexical search tree data structure storing a plurality of URLs for said received URL, said filter further operable to process said
10 received URL in response to said received URL not matching any of said plurality of URLs.

11. The system of claim 10, wherein said event notification relates to an event selected from the group consisting of a URL map event and a receive raw data event.
15

12. The system of claim 10, further comprising:
a computer readable medium with computer program logic recorded thereon for searching said lexical search tree data structure, said computer readable medium
20 comprising:

means for determining a hash value for said received URL;

means for determining a branch associated with a root node of said lexical search tree data structure corresponding to said hash value, said branch along with said root node representing at least one URL of said plurality of URLs, said branch having one or more leaf nodes linked hierarchically to one
25 another, each leaf node representing an element of said at least one URL; and

means for traversing only said branch to find a match between said at least one URL and said received URL.

13. The system of claim 12, wherein said means for determining a hash value comprises:

means for determining a first element of said received URL; and

means for determining a hash value for said first element.

5

14. The system of claim 13, wherein said means for traversing only said branch comprises means for comparing successive elements of said received URL with successive elements of said at least one URL stored in successive leaf nodes of said one or more leaf nodes so long as said successive elements of said received URL match said successive elements of said at least one URL.

10

15. The system of claim 14, wherein said means for traversing only said branch further comprises:

means for determining a twig associated with branch at a point of divergence between said at least one URL and said received URL, said twig representing a terminating substring of a second URL of said plurality of URLs; and

15

means for traversing said twig to find a match between a terminating substring of said received URL and said terminating substring represented by said twig.

20

16. The system of claim 15, wherein said means for traversing said twig comprises means for comparing successive elements of said terminating substring of said received URL with successive elements of said terminating substring of said second URL represented by said twig so long as said successive elements match.

17. A method for Uniform Resource Locator (URL) filtering, comprising:
receiving an event notification from a web server upon the occurrence of a
URL map event;

5 determining, in response to receiving said event notification, a hash value for a
URL received by said web server from a client;

determining a branch associated with a root node of a lexical search tree data
structure corresponding to said hash value, said lexical search tree data structure
storing a plurality of URLs, said branch along with said root node representing at least
one URL of said plurality of URLs, said branch having one or more leaf nodes linked
10 hierarchically to one another, each leaf node representing an element of said at least
one URL;

traversing only said branch to find a match between said received URL and
said at least one URL; and

15 processing said received URL in response to said received URL not matching
said at least one URL.

18. The system of claim 17, said receiving said event notification
comprising receiving a notification parameter from said web server, said notification
parameter pointing to a data structure storing said received URL.

20

19. The system of claim 17, further comprising notifying said web server
of a match between said received URL and said at least one URL.

20. The system of claim 17, further comprising:
25 registering with a web server to receive notification upon the occurrence of
said URL map event.

10017555-1-103101